



# **Лекция по эконометрике №1, 3 модуль**

## **Метод максимального правдоподобия**

**Демидова**

**Ольга Анатольевна**

**[https://www.hse.ru/staff/demidova\\_olga](https://www.hse.ru/staff/demidova_olga)**

**E-mail:demidova@hse.ru**

**11.01.2021**

- 1) Метод максимального правдоподобия**
- 2) Примеры применения метода максимального правдоподобия**
- 3) Свойства метода максимального правдоподобия**
- 4) Проверка гипотез с помощью теста Вальда, теста отношения правдоподобия, теста множителей Лагранжа**

**Основное предположение:**

**Известен закон распределения случайных величин, зависящий от набора параметров.**

**Оценки этих параметров подбираются таким образом, чтобы вероятность получить имеющийся набор данных была максимальной.**

$$L(\theta) = L(x_1, \dots, x_n, \theta) = P(X_1 = x_1, \dots, X_n = x_n \mid \theta) \rightarrow \max$$

**Пример 1. Найти оценку максимального правдоподобия параметра пуассоновского распределения по выборке**

**Решение:**

$$\begin{aligned} L(\lambda) &= L(x_1, \dots, x_n \mid \lambda) = P(X_1 = x_1 \mid \lambda) \cdot \dots \cdot P(X_n = x_n \mid \lambda) = \\ &= \frac{e^{-\lambda} \cdot \lambda^{x_1}}{x_1!} \cdot \dots \cdot \frac{e^{-\lambda} \cdot \lambda^{x_n}}{x_n!} = \frac{e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i}}{x_1! \dots x_n!} \end{aligned}$$

$$l(\lambda) = \ln L(\lambda) = -n\lambda + \sum_{i=1}^n x_i \cdot \ln \lambda - \ln(x_1! \dots x_n!)$$

**Необходимое условие первого порядка:**

$$l'_{\lambda}(\lambda) = -n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0$$

$$\lambda = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$$

**Достаточное условие:**

$$l''_{\lambda}(\lambda) = -\frac{\sum_{i=1}^n x_i}{\lambda^2} < 0$$

$$\hat{\lambda}_{\text{ML}} = \bar{X}$$

**Пример 2. Найти оценку максимального правдоподобия математического ожидания  $\mu$  и дисперсии  $\sigma^2$  по выборке  $x_1, \dots, x_n$**

**Решение:**

$$L(\mu, \sigma^2) = L(x_1, \dots, x_n | \mu, \sigma^2) = f_{X_1, \dots, X_n}(x_1, \dots, x_n | \mu, \sigma^2) = f_{X_1}(x_1 | \mu, \sigma^2) \cdot \dots \cdot f_{X_n}(x_n | \mu, \sigma^2) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x_1 - \mu)^2}{2\sigma^2} \right\} \cdot \dots \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x_n - \mu)^2}{2\sigma^2} \right\} =$$
$$= \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\}$$

$$l(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

## Пример 2. Необходимое условие первого порядка:

$$\frac{\partial l(\mu, \sigma^2)}{\partial \mu} = \frac{2}{2\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \Rightarrow \hat{\mu} = \bar{X}$$

$$\frac{\partial l(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \mu)^2 = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

$$\Rightarrow \hat{\mu}_{ML} = \bar{X}, \quad \hat{\sigma}_{ML}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$$

## Достаточные

условия:

$$H(\hat{\mu}_{ML}, \hat{\sigma}_{ML}^2) = \begin{pmatrix} -\frac{n}{\hat{\sigma}_{ML}^2} & 0 \\ 0 & -\frac{n}{2(\hat{\sigma}_{ML}^2)^2} \end{pmatrix}$$

# Метод максимального правдоподобия для регрессионных моделей

$$Y = X\beta + \varepsilon$$

$$\varepsilon \sim N(0_n, \sigma^2 I_n)$$

$$\begin{aligned} L(\beta, \sigma^2) &= f(\varepsilon | \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \cdot \exp\left\{-\frac{1}{2\sigma^2} \varepsilon' \varepsilon\right\} = \\ &= \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \cdot \exp\left\{-\frac{1}{2\sigma^2} (Y - X\beta)'(Y - X\beta)\right\} \end{aligned}$$

$$l(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} (Y - X\beta)'(Y - X\beta)$$

$$l(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} (Y'Y - 2\beta'X'Y + \beta'X'X\beta)$$



# Метод максимального правдоподобия для регрессионных моделей

Необходимые условия первого порядка:

$$\frac{\partial l(\beta, \sigma^2)}{\partial \beta} = -\frac{1}{2\sigma^2}(-2X'Y + 2X'X\beta) = 0$$

$$\frac{\partial l(\beta, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(Y - X\beta)'(Y - X\beta) = 0$$

$$\Rightarrow \hat{\beta} = (X'X)^{-1}X'Y, \quad \hat{\sigma}^2 = \frac{e'e}{n}, \quad e = Y - X\hat{\beta}$$

Достаточные условия:

(см. Магнус и др., 2007, 10.5)

## Общие свойства оценок МП

Если функция  $f(x)$  плотности распределения случайных величин  $X_1, \dots, X_n$  удовлетворяет некоторым условиям регулярности, то оценки МП обладают свойствами:

1) → Инвариантность. ¶

Если  $\hat{\theta}_{МП}$  — оценка максимального правдоподобия параметра  $\theta$  и  $g(\cdot)$  — непрерывная функция, то  $g(\hat{\theta})$  является оценкой максимального правдоподобия параметра  $g(\theta)$ . ¶

2) → Состоятельность (определение состоятельной оценки дано в разделе 2.3.1). ¶

3) → Асимптотическая нормальность. ¶

При  $n \rightarrow \infty$  оценка вектора параметров имеет нормальное распределение ¶

$$\hat{\theta}_{МП} \overset{a}{\sim} N(\theta, I^{-1}(\theta)), \quad \text{¶}$$

Где  $I(\theta)$  — информационная матрица Фишера, вычисляемая по формуле ¶

$$I(\theta) = -E \left[ \frac{\partial^2 l}{\partial \theta \partial \theta'} \right], \quad l — \text{функция правдоподобия}, \quad \frac{\partial^2 l}{\partial \theta \partial \theta'} — \text{матрица Гессе.} \quad \text{¶}$$

4) → Асимптотическая эффективность. ¶

Оценка ММП дисперсии каждого параметра (один из диагональных элементов ковариационной матрицы  $I^{-1}(\theta)$ ) является нижней границей для всех состоятельных асимптотически нормальных оценок этого параметра. ¶

$$Y = X\beta + \varepsilon, \text{var}[\varepsilon] = \sigma_\varepsilon^2 I$$

$$H_0 : Q\beta = q, \text{rang} Q = r$$

$$H_1 : Q\beta \neq q$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2 I) \Rightarrow$$

$$\hat{\beta}_{ML} \sim N(\beta, \sigma_\varepsilon^2 (X'X)^{-1}) \Rightarrow$$

$$Q\hat{\beta}_{ML} \sim N(Q\beta, Q \text{var}[\hat{\beta}_{ML}] Q')$$

$$W = (Q\hat{\beta} - q)' (Q \text{var}[\hat{\beta}_{ML}] Q')^{-1} (Q\hat{\beta} - q) \sim \chi^2(r)$$

# Проверка нелинейных гипотез с помощью теста Вальда

$$Y = X\beta + \varepsilon, \text{ var}[\varepsilon] = \sigma_\varepsilon^2 I$$

$$H_0: g_j(\beta) = 0, j = 1, \dots, r$$

$$H_1: \exists j: g_j(\beta) \neq 0$$

$$r = 1, g(\beta) \approx g(\hat{\beta}) + \frac{\partial g}{\partial \beta}(\hat{\beta})(\beta - \hat{\beta})$$

$$W = g'(\hat{\beta}) \left( \frac{\partial g}{\partial \beta}(\hat{\beta}) \text{ var}[\hat{\beta}_{ML}] \frac{\partial g'}{\partial \beta}(\hat{\beta}) \right)^{-1} g(\hat{\beta}) \sim \chi^2(r) = \chi^2(1)$$

# Проверка нелинейных гипотез с помощью теста Вальда

$$Y = X\beta + \varepsilon, \text{ var}[\varepsilon] = \sigma_\varepsilon^2 I$$

$$H_0: g_j(\beta) = 0, j = 1, \dots, r$$

$$H_1: \exists j: g_j(\beta) \neq 0$$

$$r > 1, g(\beta) \approx g(\hat{\beta}) + \frac{\partial g}{\partial \beta}(\hat{\beta})(\beta - \hat{\beta})$$

$$W = g'(\hat{\beta}) \left( \frac{\partial g}{\partial \beta}(\hat{\beta}) \text{ var}[\hat{\beta}_{ML}] \frac{\partial g'}{\partial \beta}(\hat{\beta}) \right)^{-1} g(\hat{\beta}) \sim \chi^2(r)$$

$$g(\beta) = \begin{pmatrix} g_1(\beta) \\ \vdots \\ g_r(\beta) \end{pmatrix}, \dots, \frac{\partial g}{\partial \beta} = \begin{pmatrix} \frac{\partial g_1}{\partial \beta_0} & \dots & \frac{\partial g_1}{\partial \beta_k} \\ \vdots & & \vdots \\ \frac{\partial g_r}{\partial \beta_0} & \dots & \frac{\partial g_r}{\partial \beta_k} \end{pmatrix}$$

**Достоинство: используются только оценки модели без ограничений**

**Недостаток: неинвариантность к способу параметризации**

# Проверка гипотез с помощью теста отношения правдоподобия

$$Y = X\beta + \varepsilon, \text{ var}[\varepsilon] = \sigma_\varepsilon^2 I$$

$$H_0: g_j(\beta) = 0, j = 1, \dots, r$$

$$H_1: \exists j: g_j(\beta) \neq 0$$

$$LR = -2(\ln L(\hat{\beta}_R) - \ln L(\hat{\beta}_{UR})) \sim \chi^2(r)$$

**В этом тесте используются оценки модели с ограничениями и без ограничений**

# Проверка нелинейных гипотез с помощью теста множителей Лагранжа

$$Y = X\beta + \varepsilon, \text{ var}[\varepsilon] = \sigma_\varepsilon^2 I$$

$$H_0: g(\beta) = \begin{pmatrix} g_1(\beta) \\ \vdots \\ g_r(\beta) \end{pmatrix} = 0,$$

$$H_1: \exists j: g_j(\beta) \neq 0, j = 1, \dots, r$$

$$H(\beta, \lambda) = l(\beta) - \lambda' g(\beta) \rightarrow \max$$

$$\frac{\partial H}{\partial \beta_k} = \frac{\partial l}{\partial \beta_k} - \lambda' \frac{\partial g}{\partial \beta_k}$$

$$\frac{\partial l(\hat{\beta}_R)}{\partial \beta_k} - \lambda' \frac{\partial g(\hat{\beta}_R)}{\partial \beta_k} = 0$$

**Если ограничения выполняются, то все множители Лагранжа равны 0, поэтому и градиент логарифмической функции правдоподобия близок к 0.**



# Проверка нелинейных гипотез с помощью теста множителей Лагранжа

$$Y = X\beta + \varepsilon, \text{ var}[\varepsilon] = \sigma_\varepsilon^2 I$$

$$H_0: g(\beta) = \begin{pmatrix} g_1(\beta) \\ \vdots \\ g_r(\beta) \end{pmatrix} = 0,$$

$$H_1: \exists j: g_j(\beta) \neq 0, j = 1, \dots, r$$

$$LM = \left( \frac{\partial l}{\partial \beta}(\hat{\beta}_R) \right)' I^{-1}(\hat{\beta}_R) \left( \frac{\partial l}{\partial \beta}(\hat{\beta}_R) \right) \sim \chi^2(r)$$

**В этом тесте используются только оценки модели с ограничениями**



NATIONAL RESEARCH  
UNIVERSITY

# Thank you for your attention!

20, Myasnitskaya str., Moscow, Russia, 101000  
Tel.: +7 (495) 628-8829, Fax: +7 (495) 628-7931  
[www.hse.ru](http://www.hse.ru)